Cue versus independent food attributes: the effect of adding attributes in choice experiments

Abstract:

We examine the effects of adding an independent food attribute on consumers' willingness to pay estimates for both cue and independent food attributes. In three separate choice experiments, a cue attribute present along the entire sequence of choices had independent food attributes enucleated and made explicit from the cue at later stages. Logit models were estimated using (1) a complete panel approach; (2) error components; and (3) utility in WTP space. Results suggest that the way a subject processes food attributes depends not only on the design dimensions but also on food attributes' functional roles. When complexity of designs increases, models that account for different sources of heterogeneity have better fit to the data.

Key words: choice experiment, choice design complexity, cue and independent food attributes; complete panel data approach, willingness to pay.

The authors would like to thank the editor, Iain Fraser, and three anonymous journal reviewers for their helpful comments and constructive suggestions. This work was partially supported by the National Research Foundation of Korea Grant #NRF-2014S1A3A2044459 and Research Council of Norway Grant # 233800.

1. Introduction

Choice experiments (CEs) have been widely employed in several fields of applied economics such as transportation, market research, health, and environmental economics, amongst others. Despite the potential scope for hypothetical bias (Lusk 2003a), its use has recently increased in consumer food choice studies, especially to investigate behavioral issues on food choice processes (Balcombe and Fraser 2011).

A key challenge in designing food CEs is how to frame experimental choice tasks in a manner that closely resembles respondents' true purchasing behavior. To reflect the increasing number of differentiated food products, CEs should feature product profiles that differ in many dimensions and attribute types. However, while the use of multiple-food attributes in choice tasks can increase choice realism, it can also complicate respondent's tasks. For this reason, practitioners typically design CEs only using a limited number of food attributes. A potentially serious weakness of this approach is that experimentally designed attributes are assumed to be independent of other omitted attributes that could be also available in the real product and relevant for consumer choice. Hence, the marginal willingness to pay (WTPs) for any attribute is implicitly considered as a value that is invariant to design dimensions. Nevertheless, if the WTP for a specific attribute depends on the number of pre-existing attributes on the product (Lusk 2003b), then the information garnered from CE studies may inaccurately reflect actual consumer purchase decisions (Lusk 2003b; Gao and Schroeder 2009), leading to biased estimates and incorrect forecasts.

In this paper, we focus on a recent and crucial debate in food CEs; i.e., the one concerning the effects of adding "independent" food attributes to choice tasks on the robustness of marginal WTP estimates of "cue" attributes (Gao and Schroeder, 2009, henceforth GS). Studying how survey respondents process cue and independent attributes has emerged as an important area of investigation because of the different functional roles of cue and independent attributes in choice behavior.

A ‘cue’ attribute (e.g., *country of origin*) is described as one that embeds in its levels some degree of information about the levels of other quality attributes not directly observed by the decision maker[1]. In other words, the levels of a ‘cue’ attribute may serve to convey information about otherwise unobservable attributes (Lusk et al. 2014)[2]. For example, the *country of origin* or even the *district of production* of a food product may be perceived by consumers as providing additional information, which might not be explicitly detailed in the product’s description (Verlegh, Steenkamp and Meulenberg 2005), perhaps due to reputation effects (Scarpa, Thiene and Marangon, 2008). Hence, with more attribute information provided about the food product, the cue attribute might lose some of its role as a proxy for overall quality. An ‘independent’ attribute, on the other hand, relates to the physical aspects of the product, whose information stands alone, irrespective of other food quality information, as it is commonly perceived to embed no further cues. For example, beef leanness is not generally associated with additional attributes of a steak, and it is hence considered an independent attribute. Thus, the value that consumers attach to an independent

[1] The information processing literature associates the word “cue” with two informational elements: the type of information examined (i.e., ‘the content’) (e.g. Jacoby, Speller and Berning 1974) and the amount of information sought (i.e., ‘the depth’) (Bettman, 1979). Hence, quality cues, also referred to as “chunks” (Simon, 1974), may provide more saliency and meaning that could then produce relative attribute dominance relations within information sets (Jacoby Olson and Haddock 1971).

[2] Hamlin (2010) also offers deeper insights into how cues are utilized and how they operate in a decision process.

attribute should be “independent” from the value attached to other attributes, especially when those attributes are not direct substitutes for it. However, during a CE study, the degree to which consumers use food attributes (both independent and cue) as quality cues might also depend on the number of attributes presented to them. For instance, given a sufficiently small set of attributes, even the "leanness of meat" might be perceived by some consumers as having a cue component for other attributes. So, a clear separation between cue and independent food attributes depends on context and is inherently subjective.

The issue about the sensitivity of CE estimates to changes in the structure of design dimensions (e.g., number and types of attributes; differences in levels, etc.) has attracted much interest. A number of studies have evaluated the effects of varying attribute information load on WTP estimates in the fields of transportation (DeShazo and Fermo 2002; Arentze et al., 2003; Hensher 2006a,b) and environmental economics (Meyerhoff, Oehlmann and Weller 2014). Results from these studies generally suggest that (i) welfare measure estimates such as WTPs are affected by the dimensionality of the experimental design, and that (ii) individuals’ processing strategies are linked not only to the dimension of CE designs but also to the functional relationship between attributes in the choice set (Hensher 2006a).

So far, only the study by GS has analyzed this issue in the context of food choice. Crucially, GS addressed the effect on the stability of WTP estimates for cue attributes when an independent attribute, previously embedded in the cue attribute, was enucleated from this and added as an explicit food descriptor. Using steak as the product of interest, the authors argued that if more information on other product attributes is provided to respondents (e.g., an attribute such as *Guaranteed Tender*), then presenting a cue attribute (e.g., *Certified U.S. Product*) may provide a weaker signal for overall product quality or for information about other attributes. To test these hypotheses, they constructed two CE surveys with different attribute numbers (one with three and four attributes, and the other one with four and five attributes). In each of the surveys, respondents were presented with a sequence of choice tasks split into two halves. GS then estimated separate choice models on data from the first and second halves of the sequences. Using the estimated marginal WTPs from a random parameter logit model, they then tested the null of no difference in WTPs between the first and second sequence for each choice experiment. Their findings suggest that the sensitivity of WTP estimates to changes in the label information was higher for attributes that are likely to provide quality cues on other missing attributes (*cue* attributes such as *Certified U.S.*) than for those which are less likely to do so (*independent* attributes e.g. *Guaranteed Tender*). They found that the marginal WTP estimate for *Certified U.S. Product* attribute decreases when the number of attributes increases from three to four, and it increases when the number of attributes increases from four to five.

Given that GS is the only study so far that has analyzed this important issue within the context of food choice, further investigation appears to be warranted to test the robustness of their crucial findings. This study extends the investigation of GS. To ease comparison across studies, we focus on the same product—steak—and a similar set of cue and independent attributes as used by GS in their first set of surveys. However, our empirical strategy differs from the one used by GS in a number of ways to tackle a number of unresolved methodological and behavioral questions related to the choice of model specification and design dimensions. One of the most formidable challenges in analyzing the effects of information load on WTP estimates due to changes in design dimensions concerns the length of the choice panel. Accordingly, we first propose an econometric approach based on the permanence of the random coefficients along the entire panel of choices by the same respondent. In this regard, we note that previous random utility coefficients analyses were conducted using separate models for the first and second part of the choice sequence, respectively, without and with the independent attribute (GS). This approach ignores the dependence between attribute coefficient values for the same individual in the two sequences. Since the respondent is the same in both sequences, the choices made by the same respondent are coherently correlated because they share idiosyncratic randomness across the utility evaluation of the attributes. By splitting the choice sequence, the information collected in the first part is not incorporated in the analysis of the second part. It is as if the process had no memory of the choices collected in the first part once it gets to the second part.

Second, we account for different sources of intra-panel variation between the two choice sequences. In this application, systematic effects associated with choices in the second half may show up as welfare effects, thereby confounding the effect of position in the sequence with the effect due to the inclusion of new independent attributes. For instance, differences in choice complexity produced by the addition of independent attributes can affect respondents' learning and fatigue (Swait and Louviere 1993; Caussade et al., 2005; Carlsson, Frykblom, and Lagerkvist 2007; Carlsson, Mørkbak and Olsen 2012; Day et al., 2012; Hess, Hensher and Daly 2012). The scale of the Gumbel error may well change between the two sequences due to other reactive factors, such as engaging in coping mechanisms used by respondents to handle the additional cognitive effort (e.g., attribute processing heuristics) (Hensher 2006a). Ignoring the possible combined and simultaneous existence of these effects of taste permanence, scale change, and coping heuristics could lead to biased parameter estimates and hence to erroneous interpretation and policy conclusions.

Finally, we estimate all our choice models with utility always specified in WTP-space (Train and Weeks 2005), rather than in preference-space and introduce an error component (EC) for every alternative different from the no-buy option to address heteroskedasticity across the buy and no-buy options.

2. Theoretical framework

Let us assume that there is a complete set of attributes $\mathbf{x}_i$ that fully describes the utility of food choice i and that the usual assumptions on the unobservable e_i and on additive utility hold. Then $U_i = \boldsymbol{\beta}'\mathbf{x}_i + e_i$. However, to avoid issues such as overloading respondents, only a sub-set $\mathbf{x}^{c}$ of the independent attributes can be used in a choice experiment. So that $\mathbf{x}$ is portioned into $\mathbf{x}^{c}$ (set of attributes included in the CEs) and $\mathbf{x}^{-c}$ (complement set of attributes excluded from the CEs) and the complete utility is $U_i = \boldsymbol{\beta}^{c\prime}\mathbf{x}^{c}_i + \boldsymbol{\beta}^{-c\prime}\mathbf{x}^{-c} + e_i$. The complement set $\mathbf{x}^{-c}$ of excluded attributes may include some for which some attributes in $\mathbf{x}^{c}$ has a "cue component". This implies that to some respondents $\mathbf{x}^{c}$ signal some values that pertain to attributes in $\mathbf{x}^{-c}$, even though these are excluded from the CE. That is, some respondents in the CE evaluate the utility as $U_i^{*} = \boldsymbol{\beta}^{c\prime}\mathbf{x}^{c}_i + \boldsymbol{\theta}'\mathbf{x}^{c}_i + e_i = [\boldsymbol{\theta}' + \boldsymbol{\beta}^{c\prime}]\,\mathbf{x}^{c}_i + e_i = \boldsymbol{\beta}^{c*\prime}\mathbf{x}^{c}_i + e_i$, where the vector $\boldsymbol{\theta}$ represents the contribution to utility that cue attributes in $\mathbf{x}^{c}$ signal with respect to the utility value of independent attributes in $\boldsymbol{\beta}^{-c\prime}\mathbf{x}^{-c}$. This makes $\boldsymbol{\beta}^{c*}$ different from the desired estimate of $\boldsymbol{\beta}^{c}$. As a consequence, the estimated marginal utility and WTP of cue attributes in $\mathbf{x}^{c}$ will also differ. As discussed earlier, food CEs can be designed using cue and independent attributes. Denote x^{a} as a cue food attribute and x^{b} as an independent attribute potentially associated with cue attribute x^{a}. The general expectation in choice modeling is that $\frac{\partial \mathrm{WTP}}{\partial x^{a}} = \frac{\partial \mathrm{WTP}}{\partial x^{a}} | x^{b}$. In other words, the marginal willingness to pay for attribute a should be invariant to the presence or absence of attribute b. If this were not the case, and a different marginal WTP is elicited for cue attributes when an independent one is specified, then the WTP estimates for the cue attributes would be contingent on information and hence would be invalid, or only conditionally valid. GS find evidence of such invalidity in their experiments.

3. Experimental procedures and survey designs

In order to test the hypothesis of a constant consumer marginal WTP for the cue attribute across varying degrees of independent attribute information, we repeat the two experiments conducted by GS (experiments A and B). But we also add a third experiment (C) to further increase the amount of attribute information offered in the CEs, which have a nested and incremental information structure.

All respondents were randomly assigned to one of three experiments. All experiments employ two CEs: CE1 that constituted the first half of the choice task sequence and CE2, which constituted the second half and included one additional independent attribute missing in CE1. Both CE1 and CE2 had 8 choice tasks, for a total sequence of 16 choices. Experiment A includes three attributes (e.g., *Certified U.S.*, *Guaranteed Tender*, and *Price*) in the first half of the sequence (A1) and four attributes in second half of sequence (e.g., *Certified U.S.*, *Guaranteed Tender*, *Guaranteed Lean*, and *Price*)(A2). Experiment B includes the same set of attributes used in A2 in the first half of the sequence (B1) and five attributes in second sequence (e.g., *Certified U.S.*, *Guaranteed Tender*, *Guaranteed Lean*, *Sell-By Date*, and *Price*). Experiment C includes the same set of attributes used in B2 in the first half of the sequence (C1) and six attributes in the second sequence (C2) (.g., *Certified U.S.*, *Guaranteed Tender*, *Guaranteed Lean*, *Sell-By Date*, *Enhanced Omega-3 fatty acids*, and *Price*). The profile of the CE studies and the attributes levels included in the experiments are reported in Table 1.

Table1. Attributes and Levels in the Choice Experiment across Experiments

Attributes (attribute levels)	**Experiment A**		**Experiment B**		**Experiment C**	
	A1	**A2**	**B1**	**B2**	**C1**	**C2**
Price ($4.64;$6.93; $9.22; $11.50)	√	√	√	√	√	√
Certified U.S. Product(absent/not absent)	√	√	√	√	√	√
Guaranteed Tender (absent/not absent)	√	√	√	√	√	√
Guaranteed Lean (absent/not absent)		√	√	√	√	√
Days before Sell-by Data (2 days; 8 days)				√	√	√
Enhanced Omega-3 fatty acids (absent/not absent)						√

Previous research indicates that experimental designs used in CE studies can significantly affect the efficiency of the final WTP estimates (Lusk and Norwood 2005). According to Scarpa, Campbell, and Hutchinson (2007) amongst others, increased estimation accuracy at given sample sizes can be achieved by adopting a sequential experimental design that progressively and iteratively optimizes some efficiency criterion. In this study, the allocation of the attribute levels was designed using a sequential experimental design with a Bayesian information structure geared to the minimization of the expected D_b-error (Scarpa Campbell and Hutchinson 2007; Ferrini and Scarpa 2007; Scarpa and Rose 2008), which is the expectation of the determinant of the asymptotic variance covariance matrix of the estimated parameters. Such expectation is computed by simulation on the basis of some prior (i.e. prior to the knowledge of the survey results) distributional assumptions. Hence, our design is developed in three sequential steps, each of which was designed to enrich the prior knowledge of such distributions. In the first step, we used as priors the estimates of a Multinomial Logit Model (MNL) from a survey conducted in 2009 to generate the two designs with 8 choice tasks each for experiment A (with three and four attributes, respectively) and for experiment B (with four and five attributes, respectively). For experiment C, the design with 5 attributes has 8 choice tasks while the design with 6 attributes requires 16 choice tasks to ensure complete identification of main effects. These are divided into two blocks of eight, each randomly assigned to respondents so as to have the same total length of the choice sequence per respondent as in the other experiments. The second step was the pilot study, which was performed in December 2013 and this provided the parameter values for the priors necessary to generate the final D_b-optimal choice design for the experiments.

Overall, (i) each design includes eight choice sets, and (ii) in each experiment, respondents are faced with 16 choice tasks, produced by combining 2 designs of 8 choice tasks each. In each choice task, respondents choose between three alternatives: two different beef steak profiles and the “no buy” option. As in GS, in order to avoid fatigue effects associated with multiple scenario valuation tasks, questions regarding respondent demographic characteristics were asked between the two halves of the sequence. Finally, in each experiment, the order of the CE questions was randomized.

4. Estimation Techniques

In our specific context, additional information about independent food attributes is made available to respondents only in the second half of the choice sequence. Hence, the panel structure of the estimator requires some adjustments. The additional attribute would explicitly address the information that might have been conjectured by some respondents as being embedded in the cue attribute in the first half of the sequence. As discussed, such an addition does not warrant separating the choice sequence into two halves and fitting two independent models to data from each half of the sequence. By the time the respondents reached the second half of the sequence, they would have achieved a certain degree of familiarity with the choice task and would have learned their tradeoffs with respect to the core set of attributes. A separate panel model fitted only to the second half of choices would not account for this effect since it would not account for the information collected on the individual distributions of taste coefficients in the first half. We posit that a more adequate formulation of the panel estimator must recognize the correlation structure of individual preferences between choices by the same respondent along the entire sequence of observed choice outcomes.

There are further considerations to make. For instance, the introduction of additional framing information is known to modify the degree of respondent's certainty in the evaluations of the utilities associated with each alternative, the so-called preference discrimination (Swait and Erden 2007). This might have an effect on the scale parameter of the Gumbel distribution, which is inversely proportional to the Gumbel error variance, inducing more determinism (discriminatory power across alternatives). The signal-to-noise ratio may therefore be modified (shifted) between the first part and the second part of the sequence. The scale of noise may be increased by making choice more stochastic (due to, for example, increased complexity of choice or fatigue) or more deterministic, hence increasing the ability of respondents to discriminate their preference due to learning (Swait and

Adamowicz 2001; DeShazo and Fermo 2002; Caussade et al., 2005; Swait and Erden 2007; Fiebig et al. 2010; Daly, Hess and Train 2012).

Next, there is cumulative evidence that utility variance differs between those alternatives that vary systematically across choice tasks due to the experimental design and those that remain the same across all choice tasks in the sequence, such as the "no-buy" option (Scarpa, Ferrini and Willis 2005; Hess and Rose 2009; Caputo, Nayga and Scarpa 2013). The former are subject to substantially higher utility variance as they are subject to new conjectures at each new choice task. Such conjecture, of course, may also involve the exact degree of embedding of independent attributes into the levels of cue attributes. An efficient way to selectively increase utility variance and induce correlation is to use a shared error component shared by the utilities of experimentally designed alternatives which involve some degree of common conjecture.

Given the above considerations, an adequate test of the effect of introducing an additional independent attribute on the marginal WTP of cue attributes can only be achieved by simultaneously addressing the following issues:

(i) adopting a complete panel approach in the random taste parameters to preserve the real panel nature of the entire sequence of food choices by the same respondent;

(ii) allowing the scale parameter of the error to be different when new independent attribute information is introduced in the label (namely in the last part of the food choice sequence in our study); and

(iii) accounting for additional covariance in the experimentally designed food profiles.

A final consideration concerns the potential lack of definition of the second central moments of the implied distribution of the ratio of two random coefficients. It is undesirable to assume a random utility structure that may imply, depending on the estimation outcomes, WTP distributions with infinite variance or implausibly “fat” tails, so as to ease inference. Random utility specified in the preference space with random attribute and cost coefficients often produces these problems in marginal WTP estimates (see Train and Weeks 2005, Scarpa, Thiene and Train 2008; Daly, Hess and Train 2012; Carson and Czajkowski 2013). While assuming a fixed cost coefficient gets around this problem, it implies a constant marginal value of money across respondents. Random utility in the (marginal) WTP-space overcomes all these shortcomings and it is undoubtedly a more appropriate approach when comparisons across treatments are made and avoids issues of scale effects present in marginal utilities (e.g., preference space). Therefore, in this study, all the models are specified in WTP-Space[3].

5. Econometric Model Specifications

[3] We also estimated choice models with utility specified in preference-space rather than WTP-space to test whether adding an independent attribute during the second half of the choice sequence causes significant effects on price coefficient estimates across all Experiments (A, B, and C). No effects were found (result are available from the authors upon request). As in Monroe (1976), this might be due to the presence of: (i) independent attribute information (e.g., *Guaranteed Tender*, etc.), (ii) no-price cue information (*Certified U.S.* label); and (iii) the no-buy option in our CE surveys.

We estimated two econometric models (i.e., Models 1 and 2 reported in the results section). The benchmark specification (i.e., Model 1 reported in the results section) is an Error Component model in WTP-space only accounting for correlation across WTPs, which represents the baseline model. The second specification (i.e., Model 2 reported in the results section) is an Error Component model in WTP-space accounting for correlation across WTPs and for both scale and marginal WTP shifters (i.e., models accounting for (i), (ii), and (iii) discussed above). Another advantage of the WTP-space framework is that it produces coefficients with a familiar and intuitive (OLS-like) interpretation for differential effects from dummy variables. These are denoted by Δ and they represent the effects on marginal WTP for attributes emerging from observed choices, after the independent attribute is included in the choice context in the second half of the sequence in each experiment, i.e., from t=9,…,t=16. The definition of the utility function for the generic steak alternative j across all experiments is as follows:

$$
\begin{aligned}
U_{jnt} = V_{jnt} + \varepsilon_{jnt} = \exp(\tau_n + \delta 1_{s2}) \times & \\
& [(\omega_{1n} + \Delta_1 1_{s2}) \mathit{US\ Certified}_{jt} + \\
& (\omega_{2n} + \Delta_2 1_{s2})\ \mathit{Tender}_{jt} + \\
& (1_{s2} \times 1_A + 1_B + 1_C)\ (\omega_{3n} + \Delta_3 1_{s2})\ \mathit{Lean}_{jt} + \\
& (1_{s2} \times 1_B + 1_C)\ (\omega_{4n} + \Delta_4 1_{s2})\ \mathit{Days\ before\ Sell\text{-}by} + \\
& (1_{s2} \times 1_C)\ (\omega_{5n} + \Delta_5 1_{s2})\ \mathit{Omega} + \\
& -\ \mathit{price}_{jt} + 1_j(\eta_{nt})] + \varepsilon_{jnt} \qquad (1)
\end{aligned}
$$

where V_{jnt} is the indirect utility function; $1_j(\cdot)$ is an indicator function that takes the value of 1 for experimentally designed food profiles; $1_{s2}(\cdot)$ is a dummy variable indicator for the second choice sequence; $1_A(\cdot)$, $1_B(\cdot)$, and $1_C(\cdot)$ are mutually exclusive dummy variables indicators for experiments A, B and C; τ_n is the common scale factor; ω_{1n} is the coefficient of the estimated WTP values; δ and Δ denote the effects of the second half of the sequence (i.e., that with the additional label information), respectively on the scale factor and on marginal WTP, and finally η_{nt} is a respondent-specific idiosyncratic error component associated only with the conjectured purchase alternatives (e.g., excluded from the no buy option).

In the above specification, the vector of random marginal WTPs for the attributes is:

$$\begin{pmatrix} \omega_{1n} \\ \vdots \\ \omega_{5n} \end{pmatrix} \sim N[\mu, \Sigma] \tag{2}$$

where the elements of Σ are to be estimated from the Cholesky matrix[4] along with the means in μ by using the maximum simulated likelihood approach and the choice data. The τ coefficients of the scale factor are also assumed to be distributed multivariate normal across respondents and are hence sub-scripted with n, while the effects on the scale factor of higher level of product information δ are fixed. Positive values of estimated δ are consistent with higher scale and hence more deterministic choice after the introduction of the independent attribute, while negative values suggest more stochastic choices (a higher noise-to-signal ratio). The exponential transformation makes the multiplicative scale/price coefficient factor strictly positive as required. The unobservable utility components denoted by ε are assumed to be i.i.d. Gumbel distributed.

In the estimation, for all the experiments (A, B, and C) and conditional on the respondent's draw of the random vector of parameters in V_n, the panel structure for the entire

[4] Cholesky matrix estimates are available upon request.

sequence of 16 choices in each of the surveys is specified to have a joint choice probability of:

$$L_n = Pr\left(y_{n1,} \cdots y_{n8,} y_{n9,} \cdots y_{n16}\right) = \prod_{t=1}^{t=16} \frac{e^{V_{jnt}}}{\sum_i e^{V_{int}}} \quad (3)$$

The unconditional distribution is simulated by using *R*=1000 Halton draws as:

$$\widetilde{L_n} = \frac{1}{R} \sum_{r=1}^{R} L_n^r \qquad (4)$$

All models are estimated using Biogeme (Bierlaire 2003) where the log of the simulated likelihood for the sample is maximized using the CFSQP algorithm (Lawrence, Zhou and Tits 1997), which is suitable for functions with several local maxima. The stability of the maximizers μ and Σ was checked using a variety of starting values. In all the EC model specifications, the price, which is treated as a continuous variable, refers to a 12-ounce steak; the rest of the qualitative attributes such as *Certified U.S., Tender, Lean, Sell-By Date,* and *Enhanced Omega-3 fatty acids* are included in the model as dummy variables. Discrete choice models are defined on utility differences. Thus, it does not matter what value is assigned to the omitted attributes. As long as they are the same across all choice alternatives, they will have no influence on choice probabilities because they imply no difference in utility. Accordingly, the omitted attributes (e.g., *Lean, Sell-buy*, and *Enhanced Omega-3 fatty acids* in the first sequence of choice of Experiments A, B, and C respectively) are, for simplicity, coded as zero.

6. Data and Results

6.1 Sample characteristics and statement of attribute attendance

A national sample of US consumers (i.e., people who have bought beef steak in the last 3 months) was randomly recruited through an email invitation by a professional market research agency (Qualtrics) and then randomly assigned to the three CE experiments (A, B, and C). A total of 201, 183, and 208 respondents completed Experiments A, B, and C, respectively. Results are reported in the supplementary materials (Table S1).

4.2 WTP-space estimates

Tables 2 reports WTP-space estimates for Experiment A, B, and C.

Table2. Estimates of EC models in WTP space of Experiment A (with three and four attributes), Experiment B (with four and five attributes), and Experiment C (with five and six attributes) (standard errors)

		Experiment A		Experiment B		Experiment C	
		Model 1	*Model 2*	*Model 1*	*Model 2*	*Model 1*	*Model 2*
WTP parameters							
No-Buy	Coeff.	-6.32***	-7.90***	-12.4***	-11.50***	-10.30***	-9.31***
		(0.80)	(1.27)	(1.93)	(1.78)	(0.54)	(0.42)
τ(n)	Mean	-0.53***	-0.59***	-0.72***	-0.51***	-0.35***	-0.51***
		(0.14)	(0.13)	(0.12)	(0.17)	(0.09)	(0.14)
	St.dev.	1.17***	1.55***	1.21***	1.18***	1.16***	1.20***
		(0.12)	(0.14)	(0.11)	(0.14)	(0.12)	(0.13)
US Certified	Mean	6.05***	8.50***	6.04***	6.44***	4.03***	4.90***
		(0.93)	(1.00)	(0.48)	(0.64)	(0.28)	(0.55)
	St.dev.	7.92***	7.32***	5.94***	5.71***	3.49***	3.67***
		(0.98)	(0.71)	(0.33)	(0.46)	(0.19)	(0.18)
Tender	Mean	3.29***	4.23***	3.40***	2.89***	1.80***	1.82***
		(0.46)	(0.52)	(0.35)	(0.31)	(0.16)	(0.28)
	St.dev.	0.93*	1.05***	1.67***	2.52***	0.24***	0.24
		(0.21)	(0.18)	(0.32)	(0.22)	(0.11)	(0.16)
Lean	Mean	2.29***	2.75***	2.09***	2.26***	1.49***	1.75***
		(0.32)	(0.36)	(0.23)	(0.28)	(0.16)	(0.23)
	St.dev.	0.33	0.12	0.21**	0.38***	0.16	0.04
		(0.25)	(0.08)	(0.11)	(0.09)	(0.11)	(0.10)

	Mean			2.19***	2.20***	1.07***	1.24***
Sell-By				(0.39)	(0.46)	(0.18)	(0.24)
				0.15	0.13	1.32***	1.29***
	St.dev.			(0.87)	(0.28)	(0.16)	(0.29)
	Mean					0.90***	1.00***
Omega						(0.24)	(0.28)
						1.64***	2.04***
	St.dev.					(0.23)	(0.31)
Error Comp.	St.dev.	6.95***	5.49***	8.40***	6.76***	6.12***	5.65***
		(0.71)	(0.72)	(1.38)	(0.76)	(0.44)	(0.58)
Scale and utility shifters							
Shift in scale (δ)			0.21*		-0.24*		0.23***
			(0.11)		(0.14)		(0.12)
Δ US certified			-1.06***		-1.04***		-1.23
			(0.11)		(0.28)		(0.24)
Δ Tender			-0.56***		1.00***		0.06
			(0.17)		(0.35)		(0.17)
Δ Lean					-0.47**		-0.28
					(0.22)		(0.18)
Δ Sell buy							-0.19
							(0.22)
Summary Statistics							
N		3216	3216	2928	2928	3328	3328
Log likelihood		-2024	-1988	-2007	-1987	-2319	-2301
AIC/N		1.271	1.251	1.389	1.378	1.415	1.407
BIC/N		1.309	1.294	1.444	1.442	1.479	1.480

N. of parameters	20	23	27	31	35	40

Note: ***, **, * indicate that parameters are statistically significant at 1%, 5% and 10% level.

As previously mentioned, two different error-component specifications in WTP-space (Models 1 and 2) are reported for each experiment. Models 1 is the basic specification accounting only for correlation across WTPs, while Models 2 adds shifts due to the introduction of the additional independent attribute in the second half of the sequence of the panel. In particular, two types of late sequence shifters are accounted for: the scale shifter denoted by δ, which accounts for net effect of learning (if positive) or fatigue (if negative), and the shifters of attribute-specific marginal WTPs, denoted by Δ. A negative and significant sign of δ is evidence of a shrinking scale—and hence a more deterministic choice often linked to relatively less cognitively complex choices—following the introduction of an independent steak attribute. A positive effect suggests a more stochastic choice, perhaps due to higher cognitive load. In contrast, a positive and significant sign of Δ indicates a WTP increase in the sequence of choices after the inclusion of the independent steak attribute; while a negative and significant sign would be consistent with the decrease of WTPs in the sequence of choices after the inclusion of the independent steak attribute. In all the models, all attribute coefficients (marginal WTPs) were specified as random, while τ_n is assumed to be log-normally distributed, but independently of the multivariate normal distribution of the marginal WTPs for beef steak attributes.

In all the models from the three experiments, the estimates of population means for the marginal WTPs are found positive and significant at the 1% level. Restrictions on the Cholesky matrix imposing preference homogeneity are strongly rejected. Finally, the distribution of error component associated with experimentally designed alternatives has a significant and large estimate for the standard deviation, indicating that utility variance is much larger for purchase than for no-purchase alternatives.

In Experiment A (first two columns of Table 2), the relative ranking based on the estimated mean of the marginal WTP distribution is consistent across all 2 Models and as follows. *Certified US* (range \$6.05 -\$8.50) has the largest value estimate, *Guaranteed Tender* (range \$3.29-\$4.23) has intermediate value estimate, and *Guaranteed Lean* (range \$2.29-\$2.75) has the lowest value. Turning to the effects on mean values of the marginal WTP estimates of an additional independent attribute (Δ) (Model 2), we find these to be negative and significant for both the *Certified U.S.* (\$ -1.06) and *Guaranteed Tender* (\$ -0.56). This finding is consistent with the findings of GS who found that when more label information was used to describe the product, the cue attribute (*Certified U.S.*) was affected more than the independent attributes (e.g., *Guaranteed Tender*). They ascribed this effect to the loss of power in terms of quality signal suffered by the cue attribute. However, in our case no statistically significant difference is found between cue and independent attributes (see 95% confidence intervals of Model 2 in Table S2 of the supplementary materials). This suggests that both cue and independent attributes are perceived by consumers as having a cue component. Finally, we provide separate evidence on the scale effects by introducing the additional independent attribute (δ) in the second half of the sequence in Model 2. The estimate for the scale shifter is positive and significant, implying more determinism in choice in the second half of the sequence and in the presence of additional food descriptors.

In Experiment B (first two columns of Table 2), the independent attribute added in the second half is *Sell-By Date*, while the independent attributes *Guaranteed Tender*, *Guaranteed Lean* and the cue attribute *Certified U.S.* were part of all choice tasks. The relative ranking of the estimated means of the marginal WTPs is as follows. *Certified U.S.* (\$6.04–\$6.44) followed by *Guaranteed Tender* (\$2.89–\$3.40), then by *Guaranteed Lean* (\$2.09–\$2.26), and by *Sell-By Date* (\$2.19–\$2.20) in both EC models, expect for the *Sell-By Date* (\$2.19) and *Guaranteed Lean* (\$2.09) attributes, which are ranked as third and fourth value estimate respectively in Model 1. Turning to the effects on the mean values of the marginal WTP estimates of an additional independent attribute (Δ) (Model 2), we find these to be negative and significant for the *Certified US* ($ -1.04) and *Guaranteed Lean* ($ -0.47), and positive and significant for *Guaranteed Tender* ($ 1.00). Finally, we report a negative and significant estimate for the scale effect (δ).

In Experiment C (last two columns of Table 2) the independent attribute added is *Enhanced Omega-3 Acids*, while all others are in the entire sequence. The relative ranking of the estimated means of the marginal WTPs for *Certified U.S.* (range $4.03–$4.90) and *Guaranteed Tender* (range $1.80–$1.82) is stable at the top, albeit with a lower value than from Experiment A and B across all error component models. Looking at the magnitude of the marginal WTP estimates for the independent attribute *Sell-By Date,* we note that this is much smaller in C than in A and B. Also, the estimated means of the marginal WTP for the additional independent attribute *Enhanced Omega-3 Acids* are smaller than those of the other independent attributes added in Experiments A *(*e.g. *Guaranteed Lean)* and B *(*e.g. *Sell-by Date)*. In our case, it seems that the diminishing marginal utility from an extra attribute is conditional to the number of attributes used as conjectured by Lusk (2003b). Unfortunately, we cannot control for the effect of order on the estimated value for this attribute coefficient (e.g., when it is in 3^{rd} or 4^{th} position). Thus, it may also be that the attribute itself has a lower value. The most interesting result emerging from Experiment C is that the WTP effects of an additional independent attribute (Δ) are found to be consistently negative and significant only for the cue attribute *Certified U.S.* ($-1.123). All other independent attributes have insignificant estimates of Δ. Hence, only the value estimates for the cue attributes are significantly affected by the addition of independent food attributes. Turning to the effects of the second half of the sequence on the scale of the error, we find these to be significant and positive (implying more deterministic choice and/or higher preference discrimination), which could be due to learning effects or better discrimination due to the additional information.

Table 2 also reports the information criteria used to decipher the relative fit of the various models. The lower the information criterion value, the better is the fit. The reduction in the AIC and BIC statistics in models 1 and 2 indicate that Model 2 is superior to Model 1 in all Experiments (e.g., A, B, and C).

7. Main findings and Conclusion

In food CEs, understanding the extent to which estimates of marginal WTPs for product or service attributes are influenced by the number and type of attributes presented to the respondents has important implications for both study design and reliability of estimates. Such implications can be extended to both hypothetical and non-hypothetical choice studies. The research agenda aims to disentangle the important relationship between value estimates and their context dependency.

To date, only the study by GS has analyzed the effect of introducing one additional food attribute in a CE on food choice. Hence, scant information is available on how WTP estimates are affected by varying the number of food attributes in a CE design; especially when information potentially embedded in cue attributes becomes explicit by the addition of independent attributes.

This study offers a novel methodological and empirical approach in analyzing the effects of adding attribute information in CEs. It builds on previous knowledge in many respects:

1) First, and most importantly, this is the first study that uses models based on a complete panel approach as opposed to an approach based on models from a split panel. This allows us to capture two different sources of intra-panel variation (differential effects) such as shifters of the scale factor and shifters of attribute-specific marginal WTP;
2) Second, the GS study and ours are the first studies to analyze the effects of food choice complexity on WTP estimates by focusing on the different information type (cue versus independent), rather than simply on the number of attributes in CE designs.

Results suggest that the number of attributes affects marginal WTP estimates, which is consistent with some previous results in transportation (Hensher 2006b,) and in environmental economics (Meyerhoff, Oehlmann and Weller 2014). They also suggest that when complexity increases in CE designs due to the addition of more attributes, changes in marginal WTP estimates not only depend on the number of attributes but also on the functional role played by the attribute type: cue or independent attributes. This finding also aligns with previous results linking individuals' processing strategies to both the functional relationship between attributes in the choice set and their number (Hensher 2006a). Most importantly, they also align with those from GS, which indicate that the WTPs for both the cue attribute (e.g., *Certified US)* and the independent attributes (*Guaranteed Tender, Guaranteed Lean, and Days Sell-by*) seem to significantly depend on the dimension of CE designs when the number of attributes is changed from 3 to 4 (i.e., Experiment A) and from 4 to 5 (i.e., Experiment B) for both cue and independent attributes. However, when the number of attributes increases from 5 to 6 (Experiment C), our results only confirm the finding of GS regarding the effects of the cue attribute on marginal WTP estimates, since no significant change is found for independent attributes. An overview of the main findings of our study is exhibited in Table 3.

Table 3. Overview of the main findings from Model 2 across Experiments

	Experiment A	Experiment B	Experiment C
Marginal WTP Rankings[1]			
US certified	1st	1st	1st
Tender	2nd	2nd	2nd
Lean	3rd	3th	3rd
Sell buy		4rd	4th
Enhanced Omega-3 fatty acid			5th
Scale and utility shifters[2]			
Shift in scale (δ)	Positive***	Negative*	Positive***
Δ US certified	Negative***	Negative***	Negative***
Δ Tender	Negative***	Positive***	Positive
Δ Lean		Negative***	Negative
Δ Sell buy			Negative

[1] Relative ranking of the marginal WTPs for the attribute information across Experiments (e.g. A, B, and C).

[2] Effects (e.g. positive and negative) and significance (e.g. ***, **, * indicate that parameters are statistically significant at 1%, 5% and 10% respectively) of the scale shifters and shifters for the attribute-specific marginal WTP across Experiments (e.g. A, B, and C).

As for the reason for this departure, we can only speculate. We suspect that the difference might be due to our use of the entire sequence of choices in the panel to estimate random coefficients. This speculation is supported by the ancillary robustness analysis we conducted in our data. In fact, when we applied the split panel approach used by GS to our data (see supplementary material – Table S3), statistically significant differences in WTP estimates do emerge for both independent and cue attributes (e.g., *Certified US, Guaranteed* T*ender, Guaranteed Lean, and Sell-By Date*). Therefore, from a methodological perspective, our study shows that the use of the entire sequence of choices in the panel, along with appropriate behavioral models, can produce different results to the ones obtained from a simple random parameter logit model using a split sequence approach. Moreover, we also show that the use of a "within subjects" approach instead of a "between subjects" approach, together with the adoption of a complete panel approach, also allows for a thorough investigation of the differential effects and shifts in behaviors across treatments in experimentally designed treatment-effect studies; such as differential information provision, mitigation of hypothetical bias, context effects, etc.

Finally, from an empirical perspective, our findings show that the functional role played by both cue and independent food attributes is affected by the dimension of the attributes space. Specifically, our CE design consists of a relatively small number of attributes (from 3 to 4 and from 4 to 5), with both independent and cue attributes exhibiting a cue component, and with a corresponding change in their marginal WTP estimate when adding a new independent attribute for both cue and independent attributes. On the other hand, with a larger attributes space (from 5 to 6), we find that only the cue attribute (*Certified US*) exhibits the cue component. It is encouraging to compare this evidence with that found by Hensher (2006a) in the field of transportation, who showed that an independent attribute such as the mean-weighted average WTP for a specific attribute (i.e., time saving), was unaffected by the design dimensionality after controlling for all design dimensions (i.e., number of choice sets, attributes, alternatives, attribute levels, and range of attribute levels).

We hope that these findings can motivate other food CE researchers to delve into this promising research area. For instance, future studies should investigate how WTPs for food attributes are affected when varying other measures of design complexity such as its entropy, the number of choice sets, attribute levels, alternatives, and ranges of attribute levels. As DeShazo and Fermo (2002:pp.141) argued: "….economists should vary complexity across survey instruments so that welfare estimates may be evaluated at either the optimal level of complexity or the level of complexity most often encountered by respondents". Further research effort should also be placed on determining whether there is symmetry in effects when increasing or decreasing attribute information load. For example, it would be interesting to know what happens to marginal WTP estimates if information on attributes is decreased from an initially richer set. That is, what if the cue and independent food attributes are first used for profile descriptions and then are removed? If a constant budget reallocation mechanism is in place, then the marginal effects on WTP for cue attributes should be positive when independent attributes for which they proxy are removed. It is also possible that the dimension of the attribute space could induce alterations in marginal WTPs through a different mechanism such as "information overload". While we recognized this in our study, we have not directly tested this effect since information overload can have broader impacts than task complexity. Future research should also examine respondents' use of "heuristics" when they intend to filter out irrelevant information when facing information overload or task complexity. Also, while it is true that each of the independent attributes may not induce a change in WTP estimates for the cue attribute, future studies should check if the joint information of multiple independent attributes could do so. Lastly, future studies should also test methodologically if a heterogeneous design such as that used by Sandor and Wedel (2005) can improve statistical efficiency.

References

Arentze, T., Borges, A., Timmermans, H. and DelMistro, R. (2003). Transport Stated Choice Responses: Effects of Task Complexity, Presentation Format and Literacy. *Transportation Research* Part E: 229-244.

Balcombe, K. and Fraser, I. (2011). A general treatment of 'don't know' responses from choice experiments. *European Review of Agricultural Economics* 38(2):171-191.

Bettman, J. R. (1979). An Information Processing Theory of Consumer Choice, Reading, Massachusetts: Addison-Wesley Publishing Company, Inc.

Bierlaire, M. 2003. *BIOGEME: a free package for the estimation of discrete choice models*, Monte Verita, Ascona, Switzerland.

Caputo, V., Nayga, R. M. Jr. and Scarpa, R. (2013). Food miles or carbon emissions? Exploring labeling preference for food transport footprint with a stated choice study. *Australian Journal of Agricultural and Resource Economics* 57:1-18.

Carlsson, F., Frykblom, P. and Lagerkvist, C.J. (2007). Consumer willingness to pay for farm animal welfare: mobile abattoirs versus transportation to slaughter. *European Review of Agricultural Economics* 34(3): 321-344.

Carlsson, F., Mørkbak, M.R., and Olsen, S.B. (2011). First Time is the Hardest: A test of ordering effects in choice experiments. *Journal of Choice Modelling* 5(2):19–37.

Carson, R.T. and Czajkowski, M. (2013). A new baseline model for estimating willingness to pay from discrete choice models. International choice modelling conference, Sydney. 2013.

Caussade, S., J., Ortuzar, de D., Rizzi, L.I. and. Hensher, D.A. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B* 39:621-640.

Daly, A.J., Hess, S. and Train, K. (2012). Assuring Finite Moments for Willingness To Pay in Random Coefficients Models. *Transportation* 39(1):19-31.

DeShazo, J.R. and Fermo, C. (2002). Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management* 44:123-143.

Ferrini, S. and Scarpa, R. (2007). Designs with a-priori information for nonmarket valuation with choice-experiments: a Monte Carlo study. *Journal of Environmental Economics and Management* 53(3):342-363.

Fiebig, D.G., Keane, M.P, Louviere, J.and Wasi N. (2010). The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. *Marketing Science* 29 (3): 393-421).

Gao, Z., and Schroeder, T.C. (2009). Effects of label information on Consumer willingness to pay. *American Journal of Agricultural Economics* 91(3):795-809.

Hamlin, R.P. (2010). Cue-based decision making. A new framework for understanding the uninvolved food consumer. *Appetit*e 55(1):89-98

Hensher, D. A. (2006)a. How Do Respondents Process Stated Choice Experiments? Attribute Consideration under Varying Information Load. J*ournal of Applied Econometrics* 21:861–78.

Hensher, D. A. (2006)b. Revealing Difference in Willingness to Pay due to the Dimensionality of Stated Choice Designs: An Initial Assessment. *Environmental and Resource Economics* 34:7–44.

Hess, S., Hensher, A.D. and Daly, A. (2012). Not Bored Yet – Revisiting Respondent Fatigue in Stated Choice Experiments. *Transportation Research Part A* 46(3):626-644.

Hess, S. and Rose, J.M. (2009). Should Reference Alternatives in Pivot Design SC Surveys Be Treated Differently? Environmental and Resource Economics 42(3): 297–317.

Jacoby, J., Olson, J.C. and Haddock, R.A. (1971). Price, brand name, and product composition characteristics as determinants of perceived quality. *Journal of Applied Psychology* 55(6):570–579.

Jacoby, J., Speller, D.E. and Berning, C.K. (1974). Brand choice behavior as a function of information load. Replication and extension. *Journal of Consumer Research* 1(1):33-423.

Lawrence, C., Zhou, J. and Tits, A. (1997). User's Guide for CFSQP Version 2.5: A C Code for Solving (Large Scale) Constrained Nonlinear (Minimax) Optimization Problems, Generating Iterates Satisfying All Inequality Constraints'(TR-94-16r1), Technical report, Institute for Systems Research, University of Maryland, College Park, MD 20742, 1997.

Lusk, J.L. (2003)a. Effects of Cheap Talk on Consumer Willingness to Pay for Golden Rice. *American Journal of Agricultural Economics* 85(4):840-856.

Lusk, J.L. (2003)b. Using Experimental Auctions for Marketing Applications: A Discusssion. *American Journal of Agricultural Economics* 35(2):349-360.

Lusk J.J., and F.B. Norwood. 2005. “Effect of Experimental Design on Choice-Based Conjoint Valuation Estimates.” *American Journal of Agricultural Economics* 87(3):771-785.

Lusk J.J., Brown, J., Mark, T., Proseku, I., Thompson, R. and Welsh, J. (2014). Consumer Behavior, Public policy, and Country-of-Origin Labeling. *Review of Agricultural Economics* 28(2):284-29.

Meyerhoff, J., Oehlmann, M. and Weller, P. (2014). The Influence of Design Dimensions on Stated Choices in an Environmental Context. *Environmental and Resource Economics* 58(2):1-23.

Monroe, K.B. (1976). The influence of price differences and brand familiarity on brans preferences. *Journal of Consumer Research* 3(1):42-49.

Sándor, Z. and Wedel M. (2005). Heterogeneous Conjoint Choice Designs. *Journal of Marketing Research* 42(2): 210-218.

Scarpa, R., Campbell, D. and Hutchinson, G. (2007). Benefit Estimates for Improvements: Sequential Bayesian Design and Respondents'Rationality in a Choice Experiment. *Land Economics* 83(4):617:634.

Scarpa, R., Ferrini, S. and Willis, K. (2005). Performance of Error Component Models for Status-Quo Effects in Choice Experiments. In R. Scarpa and A. Alberini, eds. Applications of Simulation Methods in Environmental and Resource Economics. Dordrecht, The Netherlands: Springer Publisher, pp. 247-273.

Scarpa, R. and Rose, J. M. (2008). Designs efficiency for nonmarket valuation with choice modelling: how to measure it, what to report and why. *Australian Journal of Agricultural and Resource Economics* 52:253-282.

Scarpa, R., Campbell, D. and Hutchinson, G.. (2007). Benefit Estimates for Improvements: Sequential Bayesian Design and Respondents'Rationality in a Choice Experiment. *Land Economics* 83(4):617:634.

Scarpa, R., Thiene, M., and Marangon, F. (2008). Using flexible taste distributions to value collective reputation for environmentally-friendly production methods. *Canadian Journal of Agricultural Economics* 56 (2), 145-162.

Scarpa, R.,Thiene, M. and Train, K. (2008). Utility in Willingness to Pay Space: A tool to Adress Confounding Random Scale Effects in Destination Choice to the Alp. *American Journal of Agricultural Economics* 90(4):994-1010.

Simon H.A. (1974). How big is a chunk? *Science* 183:482–488

Swait, J., and Adamowicz, W. (2001). Choice Environment, Market Complexity, and Consumer Behavior: A Theoretical and Empirical Approach for Incorporating Decision Complexity into Models of Consumer Choice. Or*ganizational Behavior and Human Decision Processes* 86(2):141-167.

Swait, J. and Erdem, T. (2007). Brand Effects On Choice and Choice Set Formation Under Uncertainty. *Marketing Science* 26(5):679-697.

Swait, J., and Louviere. J.(1993). The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing Research* 30:305-314.

Train, K., and Weeks, M. (2005). *Discrete Choice Models in Preference Space and Willing-to- Pay space*. In R. Scarpa and A. Alberini, eds. Applications of Simulation Methods in Environmental and Resource Economics. Dordrecht, The Netherlands: Springer Publisher, chapter 1, pp. 1–16.

Verlegh, P.W.J., Steenkamp, J-B. E.M. and Meulenberg, M.T.G. (2005). Country-of-Origin Effects in Consumer Processing of Advertising Claims. *International Journal of Research in Marketing* 22(2):127-139.